Anonymous Referee n°2

First of all thank you very much for your constructive review, which greatly helped to improve the clarity of the manuscript.

General comments: Parrenin presents a probabilistic model for computing multiple ice-core chronologies simultaneously. Named IceChrono v1, this model is essentially the same as the Datice model except that the optimization is done numerically rather than analytically. While this slows the computation time, it simplifies the code and makes the program accessible to more people.

IceChrono v1 is publicly available on github, a useful site that provides good version control. There are also a few additional updates of Datice with respect to allowable age constraints. The release of IceChrono v1 is potentially quite useful as this dating method can now be used by other researchers. It could also provide insight into the dating method itself as adoption of the Datice chronologies has been hampered by poor explanation of the methodology. Unfortunately, the paper suffers, as the Datice papers did before it, from mathematical descriptions that are not intuitive and from incomplete diagnostics of the resulting timescale. The IceChrono solver is shown to work, but the improvements to Datice are not tested. Evaluating this paper has been particularly difficult. Should I limit my review to just whether IceChrono v1 functions as described in the paper? Or should the paper be expected to address the limitations that have been identified with Datice and also exist with IceChrono?

We are interested in knowing the limitations of IceChrono. we do not promise to fix everything in this first version but we can at least discuss the current limitations.

The Datice chronologies (first Lemieux-Dudon et al., 2010 and then AICC2012) have not been universally accepted as the best timescales for the various ice cores, and have caused widespread timescale confusion. (A recent paper in Nature by Weber et al. 2014 is a good example of this confusion). I have structured this review in two parts: first, I focus on just the description of IceChrono and second, discuss limitations in the overall Datice/IceCrhono framework. I will let the editor decide if the manuscript should address the more general Datice/IceChrono comments. First, general comments about the manuscript:
1) Remove all discussion of the Berkner Island ice core. The underlying data for the timescale have not been published so there is no ability to evaluate the inputs to the timescale. The Berkner Island core is challenging to date due to the thin, brittle ice and the first timescale for the cores needs to be accompanied by a full description of the methods used.

We agree. This application has been removed.

2) The additional age and uncertainty constraints described in Section 3.1 (particularly the third and fourth points) do not appear to be used in either example. The utility of these additions need to be described and their implementation needs to be evaluated.

Did you mean "section 2.7" ? We now added 4 new test experiments for:

- the use of ice intervals with known durations
- the use of correlated observations

- the use of gas intervals with known durations
- the use of mix ice-gas stratigraphic links

3) Describe the mathematical equations in plain English. For instance, P6817 L2-8 discusses transforming “so-called jeffreys variables” into “Cartesian variables” but does not discuss what the benefit is. This is one example but more description is necessary throughout.

We tried to explain the different equations and their benefit in plain English:

This change of variable allows to transform Jeffreys variables into Cartesian variables (Tarantola, 2005) so as to express our problem into a least-squares problem and will allow us to reduce the number of variables to be inverted (see below).

4) The language in the paper needs significant improvement. I have not tried to edit the writing but there are many instances of basic grammatical errors. For instance, the subject and verb don’t agree in the very first sentence of the abstract: Polar ice cores provide, not provides.

The manuscript has been entirely reviewed by several people so we hope that now the language is improved.

5) The references are deficient. From the introduction, one gets the impression that only Europeans (and particularly the French) have done anything noteworthy with ice cores.

The introduction has been entirely rewritten:

Polar ice cores provide continuous records of key past features of the climate and the environment, with local, regional and global relevance (e.g., Ahmed et al., 2013; EPICA community members, 2004; NorthGRIP project members, 2004; WAIS Divide Project Members, 2013). Tracers of polar climate (e.g., Jouzel et al., 2007), ice sheet topography (NEEM community Members, 2013), water cycle (Schoenemann et al., 2014; e.g., Stenni et al., 2010; Winkler et al., 2012), aerosol deposition (e.g., Lambert et al., 2012; Wolff et al., 2006) and global atmospheric composition (e.g., Ahn and Brook, 2014; Loulergue et al., 2008; Marcott et al., 2014) measured in ice cores unveil sequences of events on seasonal to glacial-interglacial time scales.
However, the prior to interpretation of polar ice core records is the complex task of building two robust time-depth relationships, one for the tracers measured in the ice phase (e.g. water isotopes, particulates and chemical impurities) and one for those measured in the air phase (e.g. greenhouse gas concentration, isotopic composition of gases). The firn, where snow is gradually compacted into ice, constitutes the upper 50-120 m part of ice sheets. The firn is permeable and air is only locked-in at its base, at a depth level called the Lock-In Depth (LID). As a result, the entrapped air is always younger than the surrounding ice at any depth level. Through gravitational fractionation processes, LID is closely related to the isotopic composition of δ15N of N2 in air bubbles data (e.g., Buizert et al., 2015; Goujon et al., 2003; Parrenin et al., 2012a; Schwander et al., 1993). The temporal evolution of the age difference between ice and air at a given depth must therefore be estimated using firn densification modeling and air δ15N. This age difference is essential for clarifying the exact timing between changes in atmospheric CO2 concentration and Antarctic surface temperature during deglaciations (Caillon et al., 2003; e.g., Landais et al., 2013; Monnin et al., 2001; Parrenin et al., 2013; Pedro et al., 2011). However, glacial-interglacial Antarctic firn changes remain poorly understood (e.g., Capron et al., 2013).
Several strategies have been developed to build ice and gas chronologies. We briefly describe these methods, their strengths and caveats hereafter:
1) Annual layer counting (e.g., Rasmussen et al., 2006; WAIS Divide Project Members, 2013; Winstrup et al., 2012). Only applicable when accumulation rates are sufficiently high to make this

annual layer identification possible, this method provides accurate estimates of event durations and small uncertainties on the absolute age of the upper ice sections. However, the cumulative nature of the errors, associated with the increasing number of counted layers, leads to a decrease of the accuracy of absolute age with depth. For instance, the GICC05 (Greenland Ice Core Chronology 2005) composite timescale for Greenland ice cores (Rasmussen et al., 2006; Seierstad et al., 2014; Svensson et al., 2008), is associated with a maximum counting error of only 45 years at ~8.2 ka B1950 (Before 1950 C.E.). This error increases progressively with depth, reaching more than 2500 years at ~60 ka B1950. Annual layer counting techniques cannot be applied when the annual layer thickness is too small to be resolved visually, e.g. in ice cores from central East Antarctica.

2) Use of absolute age markers in ice cores. Well-dated tephra layers identified in ice cores during the last millennia provide precious constraints (e.g., Sigl et al., 2014). Beyond that period, absolute age markers are very scarce. The links between 10Be peaks and well-dated magnetic events (Raisbeck et al., 2007) have provided an age marker for the Laschamp event (Singer et al., 2009). Promising results have recently been obtained using radiochronologic dating tools (Aciego et al., 2011; Buizert et al., 2014; Dunbar et al., 2008).

3) Orbital dating in ice cores. Because there are few absolute constraints in ice cores beyond 60 ka B1950 (limit for the layer counting in the NGRIP ice core), orbital tuning is the most effective method to provide chronological constraints on ice core deepest sections. In the first orbital dating exercises, tie points were determined from the tuning of water isotopic records on insolation curves (e.g., Parrenin et al., 2004), which limits further investigations of polar climate relationships with orbital forcing. More recent chronologies tried to circumvent this assumption and focused on non-climatic orbital markers. Three complementary tracers are currently used: the δ18O of atmospheric O2 (δ18Oatm) (e.g., Bender et al., 1994; Dreyfus et al., 2007), δO2/N2 (e.g., Bender, 2002; Kawamura et al., 2007; Suwa and Bender, 2008) and the total air content (e.g., Raynaud et al., 2007). While the link between δ18Oatm and precession is explained by variations in the water cycle of the low latitudes, relationships between δO2/N2, air content and local summer insolation are understood to arise from changes in the surface snow energy budget influencing its metamorphism. Without a precise understanding of mechanisms linking these tracers to their respective orbital targets, the associated uncertainties remain large, 6 ka for δ18Oatm and 3 to 4 ka for δO2/N2 and air content (Bazin et al., 2013, 2014; Landais et al., 2012).

4) Ice core record synchronization. Inter-ice core matching exercises are undertaken to transfer absolute or orbital dating information from one ice core to another one. It generally relies on the global synchroneity of changes in atmospheric composition (CO2, CH4 concentration, and δ18Oatm) (Bender et al., 1994; Blunier and Brook, 2001; Monnin et al., 2004), the identification of volcanic sulfate spikes within a given area (Parrenin et al., 2012b; Severi et al., 2007) or the hypothesis of synchronous regional deposition of aerosols recorded as ice impurities (Seierstad et al., 2014). In the first case, limitations are associated with the smooting of atmospheric composition changes through firn air diffusion. In the second case, mismatches may arise through incorrect identification of events in different ice cores.

5) Correlation with other well-dated climatic records. In some cases, high-resolution calcite δ18O records and precise U/Th dates on speleothems have been used to adjust ice core chronologies (Barker et al., 2011; Buizert et al., 2015; Parrenin et al., 2007a). Pinning ice core and speleothem records is attractive to reduce absolute age uncertainties especially during past abrupt climatic events of glacial periods. However, these exercises rely on the assumption of simultaneous abrupt climatic changes recorded in ice core (e.g. water isotopes, CH4) and low latitudes speleothem δ18O records (mostly reflecting changes in regional atmospheric water cycle). A main limitation of this method lies in the validity of this assumption.

6) Modeling of the sedimentation process: snow accumulation, snow densification into ice, air bubbles trapping and ice flow (Goujon et al., 2003; Huybrechts et al., 2007; Johnsen et al., 2001). Glaciological modeling provides a chronology derived from the estimate of the annual layer thickness, and therefore, leads to more realistic event durations when the accumulation history and thinning function are well constrained. A side product of glaciological modelling is the *quantification of changes in surface accumulation rates, and the quantification of the initial geographical origin of ice. These additional informations are necessary to convert measurements of concentrations of chemical species in ice cores into deposition fluxes, and to correct ice core records from upstream origin effects (e.g., EPICA community members, 2006; Röthlisberger et al., 2008). Caveats are caused by unknown parameters of such glaciological models, such as amplitude of accumulation change between glacial and interglacial periods, the basal melting or the vertical velocity profile, which have a growing influence at depth.*

A common and optimal chronology for several ice cores can be built through the combination of several of these methods in the frame of a probabilistic approach. The first attempts used absolute and orbitally-tuned age markers along one ice core to constrain the unknown parameters of an ice flow model (e.g., Parrenin et al., 2001, 2004; Petit et al., 1999). This method had however several limitations. First, the uncertainties associated with the ice flow model could not be taken into account, resulting in underestimated uncertainties. Second, the stratigraphic links between ice cores were not exploited, each ice core was dated separately resulting in inconsistent chronologies. A new probabilistic approach based on a Bayesian framework was subsequently introduced. The first tool, Datice, was developed by Lemieux-Dudon et al. (2010a, 2010b). It introduced modeling errors on three canonical glaciological quantities of the dating problem: the accumulation rate, the LID of air bubbles and the vertical thinning function (i.e. ratio between the in-situ annual layer vertical thickness on its initial surface thickness). This method starts from prior (also called "background") scenario for the three glaciological parameters corresponding to an prior chronology for each ice cores. These scenarios, deduced from a modeling of the sedimentation process, are associated with an uncertainty related to the degree of confidence in these prior scenarios. A minimization approach is then applied to find the best compromise between the prior chronological information for each ice core as well as absolute and relative age markers in the ice and in the air phases. This approach has been validated through the Datice tool and applied to build the Antarctic Ice Cores Chronology 2012 (AICC2012), producing coherent ice and air timescales for five different ice cores (Bazin et al., 2013; Veres et al., 2013): EPICA Dome C (EDC), Vostok (VK), Talos Dome (TALDICE), EPICA Dronning Maul Land (EDML) and NorthGRIP (NGRIP). Further developments of Datice were performe to incorporate additional dating constraints such as the depth intervals with known durations and correlation of errors (Bazin et al., 2014). Datice provides an excellent reference for this Bayesian approach. Still, because Datice has been developed over a long term period with a continuous effort in calculation optimization through methodological improvement, the final code is difficult to access for a non-expert and cannot easily be used as a community tool. We thus identified the need for an open and user-friendly program with a performance similar to Datice but that can be more easily used and implemented by different users within the ice core community.

In this paper, we present a new probabilistic model, IceChrono_v1, based exactly on the same approach as Datice but with improvements and simplifications in the mathematical, numerical and programming aspects. We first detail the IceChrono methodology highlighting the differences to Datice (Section 2). We then perform dating experiments described in section 3 using IceChrono1. We first replicate the AICC2012 experiment, and perform 4 additional experiments to test new functionalities of IceChrono1. The results of these experiments are discussed in section 4. We summarize our main findings in the conclusions, and describe perspectives for future developments on IceChrono in section 5.

Line by line comments: Please include a table with all the variables described. There are a lot of very similar symbols and they are difficult to track down in the text.

Done.

P6812,L2 – the second sentence is too long with too many parentheses to understand

This sentence has been simplified.

P6812,L12 – I don't understand the phrase "confidence interval" and how that applies to a chronology. Do you mean the uncertainty in the chronology for each age?

This was a mistake. "confidence interval" has been changed to "standard deviation" throughout the manuscript.

P6813, L5-12 – Please rewrite this paragraph so that the ideas flow more easily.

The introduction and thus this paragraph has been fully rewritten in the revised version.

P6813, L13 – What does "these" refer to? Do you mean that there are 4 broad ways to date ice cores?

The introduction and thus this paragraph has been fully rewritten in the revised version.

P6813, L17 – support your statement that this method is generally accurate for event duration.

Sentence changed to:
"Based on the evaluation of the thickness of annuals layers, this method provides good estimates of event durations and small uncertainties on the absolute age of the shallower ice sections. "

P6815, L12 – I find equation 2 difficult to follow. Please describe in more detail.

Eq. (2) (which is now eq. (3) in the revised manuscript) is described in the text:
"The third equation means that if one virtually un-thins a depth interval between an ice depth and the synchronous air depth, one gets the Unthinned Lock-In Depth in Ice Equivalent (ULIDIE, see Figure 2)."

P6817,L1 – please describe why logarithmic correction functions are needed in plain English. I'm not sure I follow the motivation.

The associated paragraph now reads:
"The variables a_k, l_k and τ_k are always positive in our problem and generally the log-normal density of probability is more adapted for these variables, which Tarantola (2005) call Jeffreys variables. The logarithm is taken so as to transform them into Cartesian variables, with a canonical density of probability being the normal. This change of variable is necessary to describe our optimization problem as a least squares problem (see section 2.4)."

P6817,L13 – how are you solving Equation 2 for deltaDepth?

We added the following description:
"To solve Equation (3), we first integrate D_k/τ_k from the surface down to every depth in the age equation grid, i.e. we have a correspondence table between real depths and unthinned-ice-equivalent (UIE) depths. Then for every air real depth in the age equation grid, we obtain the air UIE depth from the correspondence table. Then we subtract the second member of Equation (3) to this air UIE depth to get the ice UIE depth. Then we use the correspondence table to obtain the ice real depth and the Δdepth."

P6821,L15 – I don't understand "A class exist for the ice core object and does:"
P6821,L14 – This paragraph could use more explanation about object oriented

paradigms and object classes

We now introduce the main principle of the object oriented paradigm and we give more details on the object oriented structure of IceChrono v1:

The core of the code is about 1000 lines long (including white lines and comments) and is built using an object oriented paradigm. In such an object oriented language, apart for the classical type of variables (integer, real, characters, etc.), one can define his own classes of objects, containing variables and functions. In IceChrono1, a class exists for the ice core object. It contains the variables related to this ice core: the age equation grid, the correction function grids, the prior scenarios and their associated standard deviations and correlation matrices, the relative density profile, the correction functions, the observations and their associated standard deviations and correlation matrices and the resulting calculated variables (accumulation, LID and thinning, ice and air ages, Δdepth, ice and air layer thickness, etc.). It also contains functions performing the following tasks: the initialization of the ice core (i.e. reading of the parameters, priors and observations), the calculation of the age model, the calculation of the residuals, the calculation of the forward model Jacobian, the calculation of the standard deviations, the construction of the figures (for ice age, air age, accumulation, LID, thinning, ice layer thickness and Δdepth) and the saving of the results. A class also exists for each ice core pair. It contains all the stratigraphic links and their associated standard deviation and correlation matrices relative to this ice core pair. It also contains functions that perform the following tasks: the initialization of the ice core pair, the calculation of the residuals, the construction of the figures (for ice-ice links, air-air links, ice-air links and air-ice links). The main program is kept as simple and straightforward as possible.

P6821,L25 – This paragraph needs better explanation. I can’t follow what is being done “inside each term of the cost function” and why a “change of variable” is occurring. Please write out the steps and add a plain English explanation.

We wrote the steps using equations so we hope it is now clear:

We used the LM algorithm as implemented in the 'leastsq' function from the 'scipy.optimize' library, which also provides an automatic convergence criteria. It does not try to minimize directly the cost function, but rather a residual vector, the components of this residual vector being supposed independent from each others and with a unit standard deviation. Inside each term of the cost function:

$$\boldsymbol{R}^{\mathrm{T}}\mathbf{P}^{-1}\boldsymbol{R}\,, \tag{1}$$

We allow defining a correlation matrix ***P*** *so that the residuals can actually be correlated. We thus used a Cholesky decomposition of* ***P****:*

$$\mathbf{P}=\mathbf{P}^{1/2}\mathbf{P}^{1/2}\,, \tag{2}$$

and a change of variable:

$$\boldsymbol{R}'=\left(\mathbf{P}^{1/2}\right)^{-1}\boldsymbol{R}\,, \tag{3}$$

to transform, the residual vector into a vector composed of independent variables with unit standard deviation. The associated term of the cost function can now be written:

$$\left(\boldsymbol{R}'\right)^{T}\boldsymbol{R}'\,, \tag{4}$$

that is, the residuals are now independent and with a unit standard deviation.

P6822,L8 – how does this Datice assumption really differ from the IceChrono assumption. In describing equation 2, the author assumes a constant value of 1 for thinning in the firn. So isn’t the Datice assumption of a constant thinning function at depth very similar? I’m guessing there is a subtlety here that I’m missing and should be better explained. It also seems like in the AICC2012 comparison later in the paper, the effect of this difference should be specifically diagnosed.

We do not assume that thinning is 1 in the firn. We assume that thinning in the firn at the time of deposition was the same than at present. This is different from the Datice assumption, which assume that thinning between z-Δdepth and z is constant.

P2822,L14 – Please explain why adding mixed ice-gas and gas-ice stratigraphic links is important. You comment that this is new with respect to Datice, but never explain (1) whether the lack of this functionality limits AICC2012, (2) if you used any ice-air or air-ice links in the IceChrono version of AICC-2012, or (3) and concrete example of ice-air and air-ice links.

We added the following sentence to give an example of mix links:
"A concrete example of the use of mix ice-gas stratigraphic links could be the synchronization of Dansgaard-Oechger events recorded in the methane records in Antarctica and in the ice isotope records in Greenland. This would be especially useful if the methane records in Greenland are not yet available."
We also added a test experiment which uses mix ice-air stratigraphic links.

P2822,L17 – The allowance of correlated errors gets considerable attention in this manuscript, yet I get the impression at both the AICC-like and Berkner Island test cases don't make use of this functionality. The writing even seems to emphasize that you don't need to input error correlation for IceChrono to run. If this is an important advance, the impact of it needs to be assessed.

We now present a test experiment on the NGRIP core for dated ice intervals with correlation.
For convenience, the prescription of the correlation matrix in IceChrono is indeed optional and by default, it is assumed equal to the identity matrix.

P2822,L20 – Using a numerical solver of the residuals means that there is the potential to find local minima in the cost function. The paper needs to test whether the solver is robust to different initial conditions.

We tried the solver with 5 different initial conditions taken according to the prior and we checked that they all converge toward the same solution.

P6823,L4 – I was able to find and access IceChrono easily. I did not download and compile it. My quick impression is that it could use more documentation.

The documentation has been reviewed and improved. If you have any other specific request regarding the documentation, please feel free to send it to us.
(Note that IceChrono does not need to be compiled since python is an interpreted language.)

P6823,L11 – I think calling AICC2012 "the last official chronology" is a little strong. AICC2012 is "an official" chronology, though not universally accepted, and I doubt it will be "the last" official chronology.

Corrected to "most recent".

P6823,L14 – I'm confused about correlation matrices. Here you say the Datice background correlation matrices have a Gaussian shape. But on P6822,L18 you write that Datice only allows correlated errors for dated ice intervals. Please explain clearly in the text.

The most recent version of Datice actually allows defining correlation matrices for all types of observations. This sentence has therefore been corrected.

P6823,L17 – Please diagnose and describe the effect of changing from Gaussian to triangular matrices. What is the resulting difference in the timescale and uncertainty from making this change alone? What is the change in computation time?

The problem of gaussian-shape correlation matrices is not a computation time problem. The usual solver to perform a Cholesky decomposition simply do not work for gaussian-shape matrices, because they are too close to singular matrices. Instead of spending a lot of time to solve this numerical problem just for the sake of comparison with Datice/AICC2012, we decided to use triangular-shape matrices.

P6823,L11 – More description is needed of the AICC2012-like inputs. Please describe at least all the different types of dating information used and provide specific references to the AICC2012 work (i.e. if you are using the same gas tie points, reference the appropriate tables in Veres et al. 2013 and Bazin et al. 2013). There is new functionality in IceChrono, is any of this employed? The reader should not be expected to have read all the Datice papers in detail.

We now describe in more details the AICC2012 dating information used and make a reference to the supplementary material of Bazin et al. (2013) and Veres et al. (2013) where all the data are available:

IceChrono1 is similar in scope to the Datice model (Lemieux-Dudon et al., 2010a, 2010b). Datice has been used to build AICC2012, the most recent official chronology for the 4 Antarctic ice cores EPICA Dome C (EDC), Vostok (VK), Talos Dome (TALDICE), EPICA Dronning Maud Land (EDML) and the Greenland ice core NorthGRIP (NGRIP) (Bazin et al., 2013; Veres et al., 2013). The AICC2012 experiment was based on a previous experiment (Lemieux-Dudon et al., 2010a) on 3 Antarctic ice cores (EDC, VK, EDML) and one Greenland ice core (NGRIP), but with updated chronological information. All chronological informations are available in the supplementary material of Bazin et al. (2013) and Veres et al. (2013). This experiment integrates orbital tuning constraints based on the $\delta^{18}O_{atm}$, O_2/N_2 and air content records (Bazin et al., 2013, p.201; Dreyfus et al., 2007, 2008; Landais et al., 2012; Lipenkov et al., 2011; Raynaud et al., 2007; Suwa and Bender, 2008), layer counting on NGRIP back to 60 kyr B1950 (Svensson et al., 2008 and references theirein), a tephra layer (Narcisi et al., 2006) dated independently at 93.2±4.4 kyr B1950 (Dunbar et al., 2008), the Laschamp geomagnetic excursion at 40.65±0.95 kyr B1950 (Singer et al., 2009) and the Brunhes-Matuyama geomagnetic reversal at ~780.3±10 kyr B1950 and its precursor at 798.3±10 kyr B1950 (Dreyfus et al., 2008) identified in the ^{10}Be records (Raisbeck et al., 2006, 2007; Yiou et al., 1997), a Holocene ^{10}Be-dendrochronology tie point on Vostok at 7.18±0.1 kyr B1950 (Bard et al., 1997; Raisbeck et al., 1998), Δdepth observations at NGRIP obtained by comparing the $\delta^{18}O_{ice}$ and $\delta^{15}N$ records (Capron et al., 2010; Huber et al., 2006; Landais et al., 2004, 2005), air synchronization tie points using the CH_4 records (Buiron et al., 2011; Capron et al., 2010; Landais et al., 2006; Lemieux-Dudon et al., 2010a; Loulergue et al., 2007; Schilt et al., 2010; Schüpbach et al., 2011) *and the $\delta^{18}O$ records (Bazin et al., 2013) and ice synchronization tie points using the volcanic records (Parrenin et al., 2012b; Severi et al., 2007, 2012; Svensson et al., 2013; Udisti et al., 2004; Vinther et al., 2013) and the ^{10}Be records (Raisbeck et al., 2007).*

P6824 – as discussed above, this section is lacking sufficient detail in the comparison with AICC2012.

See above.

P6824, L9 – Please provide more description of the different uncertainties compared to AICC2012.

Why does Datice have such a larger uncertainty at the Laschamp event if both methods are using the same inputs?

For the AICC2012-like experiment, we now compare directly with the output of Datice and not with the AICC2012 stated uncertainty. The abnormally large uncertainty of AICC2012 at the Laschamp event is actually not present in the output of Datice.

P6824,L22 – remove this entire section. The first dating of the deep Berkner Island core needs its own paper with the data published. The records are not straightforward and will require significant explanation.

The application on the Berkner Island is now removed in the revised manuscript.

P6827,L10 – adding in glaciological models is a great next step. I think this paper would benefit from an extended discussion of why this is a limitation of IceChrono.

We added the sentence:
“Indeed, sedimentation models also have their own poorly-known parameters and it is why they need to be integrated into the optimization process.”

P6827,L18 – The appendix seems both repetitive and under-explained. All of the odd-numbered equations appear to be the same except that they are for different classes of age markers. I think the appendix would be improved by condensing the odd equations and providing explanation (not just a description of the variables) for the even equations.

We agree this appendix is a bit repetitive. The basic principle of all the terms is explained in equations (18) and (19). Now during the submission process, the editor reckoned that a detailed description of all the terms individually would help to better understand, hence this appendix.

Figure 2 – This figure would benefit from a second panel with the most recent 60 ka (when the chronologies are tied to the GICC05 annual chronology). The ice and gas chronologies as well as the ice and gas uncertainties should be shown.

We now show both the ice and gas chronologies and we added figure 5 which is a zoom on the last 60 kyr.

Figures – The removal of the Berkner Island figures will allow many more figures diagnosing the differences of the Datice/IceChrono methologies and the AICC example.

The Berkner example has been removed. The AICC2012-VHR experiment and its results are now provided as a supplementary material. This will allow the readers to analyze and modify this experiment. We did not include all the figures of this experiment in the main text since there are so many of them (7 figures per drilling and 4 figures per drilling couple!).

This second part of the review describes some of the limitations of the Datice/IceChrono methodology. While I will let the editor decide if this manuscript should address these issues, I want to emphasize that the manuscript would be much improved by taking a more comprehensive view. If the goal is to get IceChrono accepted as the best way to date ice cores, then a full description and real evaluation of the resulting chronologies will go a long way to achieving this goal.
The Datice and IceChrono methodologies have recognized limitations, which is part of why the chronologies have not been universally adopted. But there are two other reasons as well:

1) The methods have never been well explained. The Lemieux-Dudon papers are nearly impenetrable and the AICC2012 papers add very little in methodological description.
2) The methodologies have never been shown to actually yield improved timescales. Yes, Datice and IceChrono can produce timescales for all the various ice cores, but there has never been a test case that truly evaluates the resulting timescales. A synthetic test case has never been performed; why not develop "known" timescales, add noise, employ Datice/IceChrono, and then compare the inferred timescales to the "known" timescales? This would allow a much improved understanding of the methodology and greatly improve the confidence in the inferred timescales.
This work by uses the agreement with AICC2012 as validation for IceChrono. I admit that I bring a bias against the Datice chronologies because there are too many oddities: -the small glacial delta-ages of EDML -the reversals in the thinning function of EDML and Talos Dome in glacial-transition ice -the same uncertainties for the ice and gas timescales that have not been explained.

We think we should separate two different kinds of critics regarding AICC2012/DatIce: the critics toward the DatIce method (which aims at bringing different kind of chronological information together) and the critics toward the AICC2012 experiment, which uses a particular set of chronological information.
Here we developed a probabilistic dating method, comparable to DatIce in its scope, but which fixes some of its issues. So we are ready to discuss the methodological aspects (the model), but not the AICC2012 experiment in particular (the simulation).

I worry IceChrono may perpetuate many of the Datice methodological problems and create more, not less, confusion about ice-core timescales.
A few of the issues identified with Datice:
A) Ice-ice stratigraphic links are predominantly matches in sulfur peaks between cores. These matches are either correct, or wrong. However, the Datice/IceChrono methology assigns an uncertainty (I think Gaussian) to them. The final timescale results in the links no longer being exact, invalidating the premise of the links in the first place.
This limitation needs to described with guidance about what types of analysis the Datice/IceChrono chronologies are not appropriate for.

This is clearly a limitation of the IceChrono (and DatIce) model in its current state.
All uncertainties are assumed gaussian and we agree that the gaussian model does not represent correctly our state of knowledge regarding the volcanic peaks matching. So for now, we assume that the volcanic matching is perfectly done by the operator within the width of the volcanic peak.
We added a section in the discussion regarding the limitations of IceChrono v1.
We also mention in the perspective the coupling with the volcanic matching problem:

1) All uncertainties are assumed Gaussian. While the Gaussian probability density functions are the most often encountered in Science, it is not always appropriate. For example, the volcanic synchronization of ice cores (e.g., Parrenin et al., 2012b) is ambiguous: a volcanic peak in one ice core can sometimes be synchronized to several other volcanic peaks in the other ice core. IceChrono1 therefore assumes that the features recognition has been performed without ambiguity and that the only uncertainty remaining is the exact match of the two features within their durations.

B) The only glaciological constraints are in background scenarios. This leads to chronologies that are not glaciologically consistent. For the inferred thinning functions, this is revealed through reversals where deeper, older ice has thinned less than shallower, younger ice. For most ice core sites, this is not physically realistic (compressive flow can allow these reversals, but this is not appropriate for most dome sites). Impurity concentrations have been suggested as the cause, but these ideas ignore continuity and are implausible.

We disagree that a reversal in thinning is not glaciologically consistent. For example, Parrenin et al. (JGR, 2004) show that this kind of reversal occurs for the Vostok ice core, due to a thinner ice thickness upstream of the drilling site. We agree this is more difficult to explain for a dome site, but one is never sure that a dome has stay stable during the past and some anomalous flow can also occur (Dreyfus et al., CP, 2007).
The idea of the IceChrono (or DatIce) model was that ice flow models are imperfect, and that we wanted to use their information, but in a weak way, not in a strong way.

C) The gas-age and ice-age uncertainties are the same. This is odd because there should be different uncertainties based on whether the age markers are in the ice or gas. In tracking down this oddity, I found this was instituted after a reviewer of the Bazin et al. (2013) AICC paper found that the ice age uncertainties were smaller even when the age markers were in the gas phase and the gas age uncertainty should have been smaller. The authors decided to make the uncertainty for both the ice and gas equal to the larger of the two uncertainties. This approach concerns me because (1) the error in the Datice methodology was never diagnosed and (2) the uncertainties are incorrect. While this may seem a minor point, it shows clearly that the output of Datice is not fully understood. If Datice cannot do a simple uncertainty correctly, why should anyone have faith that the more complicated implementations find an “optimal” chronology. This work has not provided enough detail to diagnose whether IceChrono suffers from the same problem.

There is a known limitation in the way the confidence intervals were calculated in AICC2012 using Datice. For example, the air age uncertainty at EDC is too large during the last glacial part with respect to the observations which have been used.
As far as we have tested, this problem does not exist in IceChrono:

During the last glacial period, there are many CH_4 Antarctica-NGRIP stratigraphic links with uncertainties of a few centuries. The NGRIP chronology being tightly constrained to GICC05 within 50 yr, it is expected that the gas age uncertainty at EDC sometimes decreases below 1000 yr during this time period. The posterior uncertainty calculated by IceChrono is therefore consistent with the chronological information used, in contradiction to that calculated by Datice.

While possibly outside the scope of this manuscript, addressing these (and other concerns) of the Datice/IceChrono methodology would provide confidence in the resulting chronologies.